考研数学考点频率及03-19年平均分

1. 数据来源

通过百度搜索相关关键词，得到完整的原始数据复制粘贴或手动输入到excel中形成数据文件。

1. 数据可视化过程描述

用到的知识点：Matplotlib精细绘图、

柱状图、

参数调整、

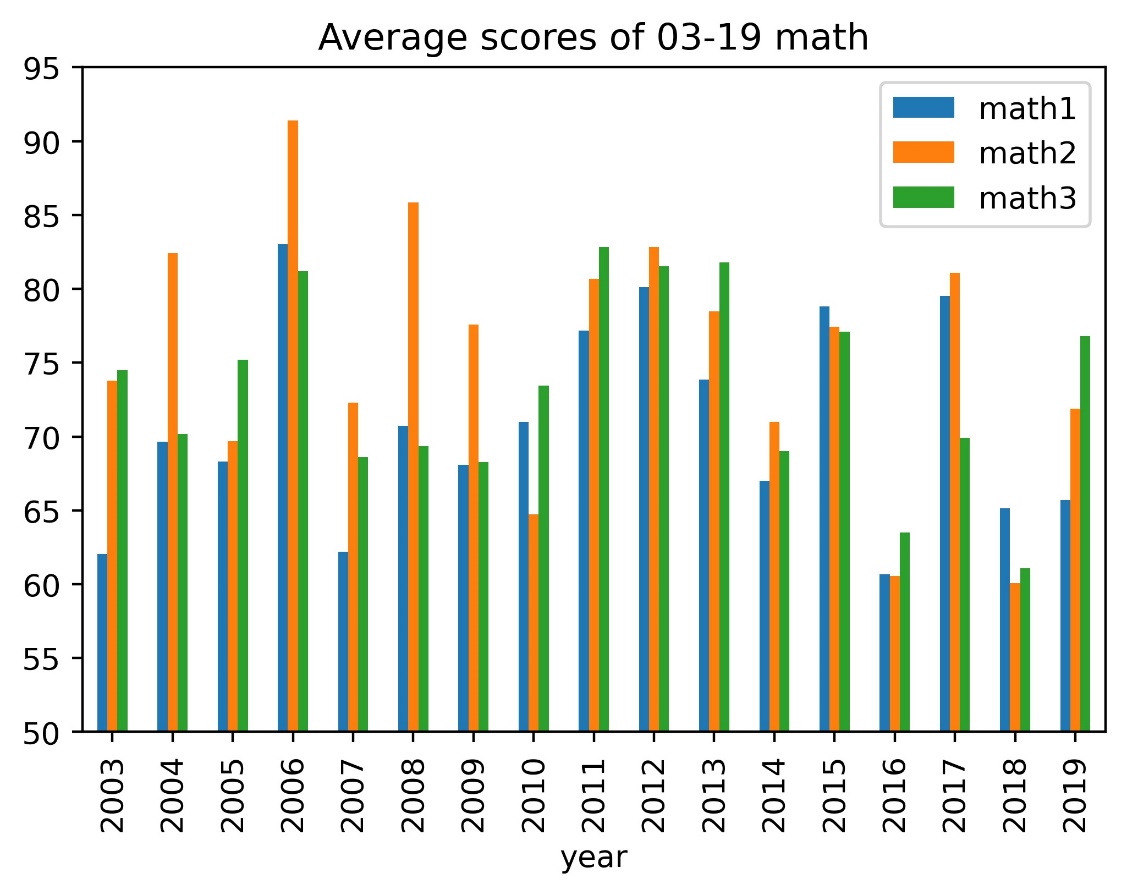
保存高清图片

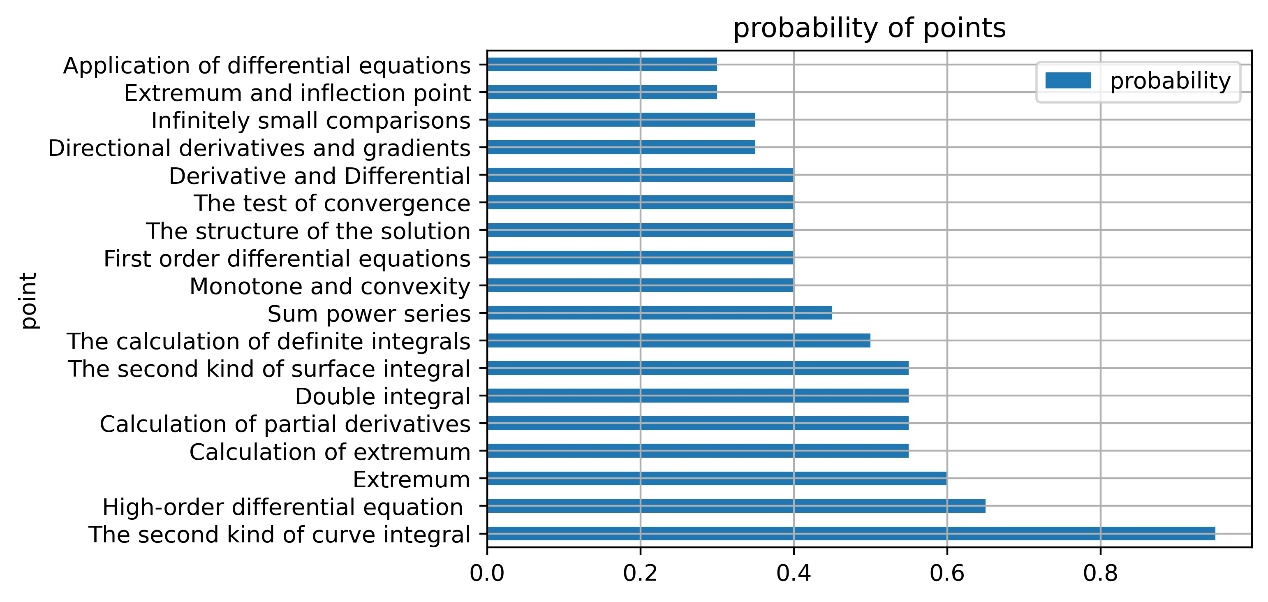
遇到的困难：

1. x坐标字数太多。
2. 原数据是中文，无法在图形中显示。
3. 数据柱过长，无法直观体现数据在不同年份的差异。

解决方法：

1. 原计划打算使用折线图，经过查找网上资料后发现可能的解决方法之一是使用水平柱状图，故改成水平柱状图。
2. 通过将中文换成等意的英文，使可视化图形上可以显示文字信息。
3. 调整Y轴参数，从而加大柱状图纵向的数据差异。
4. 图形展示





1. 代码展示

import pandas as pd

import matplotlib.pyplot as plt

data=pd.read\_excel('c:\\data\\score.xlsx',index\_col='year')

data1=data[['math1','math2','math3']]

data1.plot(kind='bar',title='Average scores of 03-19 math',

ylim=[50,95])

plt.savefig('score.jpg',dpi=400,bbox\_inches='tight')

data=pd.read\_excel('c:\\data\\point.xlsx',index\_col='point')

data.plot(kind='barh',title='probability of points',

grid=True,use\_index=True)

plt.savefig('point.jpg',dpi=400,bbox\_inches='tight')